LIBRARY
TECHNICAL REPORT SECTION
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93940

NPS-55Hh75102

NAVAL POSTGRADUATE SCHOOL
Monterey, California

A BRANCH AND BOUND METHOD FOR

NONSEPARABLE NONCONVEX OPTIMIZATION

by

James K. Hartman

October 1975

Approved for public release; distribution unlimited

Prepared for:
Office of Naval Research, Arlington, Virginia 22217

FEDDOCS
D 208.14/2:NPS-55Hh75102

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral Isham Linder
Superintendent

Jack R. Borsting
Provost

The work reported herein was supported by the Foundation Research Program of the Naval Postgraduate School with funds provided by the Chief of Naval Research.

Reproduction of all or part of this report is authorized.

This report was prepared by:

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS-55Hh75102	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A BRANCH AND BOUND METHOD FOR NONSEPARABLE NONCONVEX OPTIMIZATION		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) James K. Hartman		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940 Code 55Hh		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61152N,RR 000-01-10 N0001476WR60052
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research, Arlington, Va.		12. REPORT DATE October 1975 13. NUMBER OF PAGES 11
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear programming, nonconvex, optimization, branch and bound.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper a nonconvex programming algorithm which was developed originally for separable programming problems is formally extended to apply to nonseparable problems also. It is shown that the basic steps of the method can be modified so that separability is no restriction.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A. Background

In [1] a method was developed for solving separable nonconvex optimization problems. The method can be viewed as an extension of the Dantzig Wolfe convex programming algorithm to nonconvex problems. The extension involves imbedding the Dantzig Wolfe approximating linear programs in a branch and bound algorithm. At each stage of the branch and bound search a restricted master approximating linear program is solved over a subset of the original feasible region. Dual variables from the L.P. solution are used in (nonconvex) single variable Lagrangian subproblem minimizations which price out new trial solutions for the master L.P. The results of the Lagrangian subproblem solutions are

a) a global optimality test

b) (perhaps) new columns for the master L.P.

c) (perhaps) a branch in the branch and bound algorithm to further partition the feasible region.

Details of the method are given in [1], and an extension to ε-optimality is discussed in [2].

The purpose of the paper is to show that essentially the same method can be applied to nonseparable nonconvex optimization problems. The changes involve

a) the Lagrangian subproblem minimizations no longer decompose into single variable minimizations.

b) the branching rules for branch and bound are modified. The new rules are a significant improvement and would probably be advantageous in the separable case also.

The primary advantage of the original method, which is retained in the extension of this paper, is that nonconvexity only needs to be directly

considered in the essentially unconstrained Lagrangian subproblems. The efficiency of various methods for such subproblems has been considered in [3].

The extension is "formal" in the sense that although all the mathematical operations are valid and capable of routine computer implementation, the convergence of the algorithm for arbitrary nonconvex problems has not yet been demonstrated.

The report is intended to be read along with [1] although some material presented there has been restated here in nonseparated form for clarity.

B. Preliminary Results

We consider the general possibly nonconvex bounded nonlinear optimization problem

$$\begin{aligned} \text{NLP} \quad & \min f(x) \\ \text{subject to} \quad & g_i(x) \le 0 \qquad i=1,\dots, m \\ & a_j \le x_j \le b_j \quad j=1,\dots, n. \end{aligned} \tag{1}$$

$$\text{Let} \quad C = \{x | a_j \le x_j \le b_j \quad \forall_j\} . \tag{2}$$

Consider the linearization of NLP obtained by selecting vectors $x_k \varepsilon C$, k=1, 2, ..., r, and defining for each x_k a convex combination weight variable $\lambda_k \ge 0$ with $\sum_k \lambda_k = 1$. Let P_λ be the restricted master approximating linear program defined as

$$\begin{aligned} P_\lambda \qquad & \min \sum_{k=1}^{r} \lambda_k f(x_k) \\ & \text{subject to} \sum_{k=1}^{r} \lambda_k g_i(x_k) \le 0 \qquad i=1,\dots, m \\ & \sum_{k=1}^{r} \lambda_k = 1 \\ & \lambda_k \ge 0 \qquad k=1,\dots, r \end{aligned} \tag{3}$$

If λ^* is the primal optimal solution to the linear program, then

$$x^* = \sum_k \lambda_k^* x_k \tag{4}$$

defines an approximate solution point for the original NLP.

If NLP is a convex program, then the relationships between NLP and P_λ are well understood. These relationships have been used to develop a column generation algorithm which uses the dual variables from P_λ in Lagrangian subproblems to iteratively find new vectors x_k and hence new columns for the restricted master P_λ. This method, which can be viewed as the nonlinear analogue of the Dantzig Wolfe Decomposition Principle has been proved to converge in [4].

In the nonconvex case the P_λ and NLP relationships are not as simply understood. In particular, P_λ sometimes underestimates and sometimes overestimates the original NLP. Nevertheless the P_λ linearization can be used in an algorithm for solving NLP.

Suppose that $\pi \epsilon R^m$, $\sigma \epsilon R^1$ are the dual variables for P_λ, and define the Lagrangian function for NLP

$$L(x,\pi) = f(x) - \sum_{i=1}^{m} \pi_i g_i(x) \ . \tag{5}$$

The dual problem to P_λ is

$$\begin{aligned} \text{Dual} \quad & \max \sigma \\ & \text{subject to } \sum_i \pi_i g_i(x_k) + \sigma \le f(x_k) \\ & \qquad \pi \le 0 \text{ , } \sigma \text{ unrestricted} \end{aligned} \tag{6}$$

which can easily be rewritten as

$$\max_{\pi \le 0} \quad \min_{k=1,\dots,r} \quad L(x_k,\pi) \ . \tag{7}$$

If the inner minimization were over all $x \varepsilon C$ then (7) would be the standard Lagrangian dual for NLP.

<u>Lemma</u> 1 (Lower bound)

x^* feasible for NLP, $\pi \le 0 \Rightarrow$

$$\min_{x \varepsilon C} L(x,\pi) \le f(x^*) . \tag{8}$$

<u>Proof</u>

$$\min_{x \varepsilon C} L(x,\pi) \le L(x^*,\pi) = f(x^*) - \sum_i \pi_i g_i(x^*) \le f(x^*)$$

since $\pi_i \le 0$ and $g_i(x^*) \le 0$.□.

Suppose we solve P_λ obtaining optimal primal variables λ and optimal dual variables π,σ with objective function value Z $(=\sigma)$. Let $\hat{x}$ globally solve the nonconvex Lagrangian subproblem

$$\min_{x \varepsilon C} L(x,\pi) , \tag{9}$$

and let $x^* = \sum_k \lambda_k x_k$ as in (4).

<u>Theorem 1</u> (Optimality test)

If a) $f(x^*) \le Z$

b) $g_i(x^*) \le 0$

c) $L(\hat{x},\pi) \ge \sigma$

then x^* is globally optimal for NLP.

<u>Proof</u>

$$Z = \sigma \le L(\hat{x},\pi) \le \min\{f(x) \mid \underset{\text{for NLP}}{x \text{ feasible}}\} \le f(x^*) \le Z$$

Thus x^* solves NLP . .□.

A version of this theorem which allows for tolerances in conditions a), b), c), and the Lagrangian minimization and which implies ε-global optimality can be easily developed as in [2] .

The primary value of Theorem 1 is that if optimality is not achieved, then it suggests further actions for the optimization algorithm. In particular, if c) is violated, then the vector $\hat{x}$ generates a column $[f(\hat{x}), g(\hat{x}), 1]$ with negative reduced cost in the simplex tableau for P_λ . Thus $\hat{x}$ should be incorporated as a new x_k point. If a) or b) is violated, then the NLP is not convex at the current convex combination x^* . In this case the algorithm must resort to branch and bound to enforce a different convex combination which (hopefully) does not violate convexity so badly. In [1] the choice of variable on which to branch was simple and depended on the separated components f_j , g_{ij} of the separable functions f and g_i, where for example,

$$f(x) = \sum_{j=1}^{n} f_j(x_j) \tag{10}$$

The major point of the paper is that reasonable selection rules can be developed even in nonseparable cases. Before developing these rules we state the entire algorithm for the nonseparable case in detail.

C. The Algorithm

Step 1 Initialization

Choose an initial set of vectors $x_k \in C$. Let P_λ^t with $t=1$ (= subproblem counter) be the P_λ program corresponding to this initial set. Let $C_j^t = [a_j, b_j]$. Let $L^t = -\infty$ be the current largest lower bound for P_λ^t . Let $F^o = +\infty$ be the value of $f(x)$ for the best incumbent feasible solution to NLP found so far. Place P_λ^1 on a list of subproblems and go to step 2.

Step 2 Linear Program

If the list of subproblems is empty, stop--the incumbent solution is global optimal. Otherwise select a problem P_λ^t from the list (see

discussion in section D.) and solve it yielding optimal value Z with optimal primal variables λ and optimal dual variables π,σ. (If P_λ^t is infeasible then the method is slightly modified. See [1] for details.) Go to step 3.

Step 3 Lagrangian Minimization

Solve the nonconvex Lagrangian problem $\min_{x \varepsilon C} L(x,\pi)$ giving solution $\hat{x}$. Let $B = L(\hat{x},\pi)$. If $B \geq F^o$ then fathom P_λ^t and go to step 2. If $F^o > B > L^t$ then increase the value of the bound for P_λ to $L^t = B$ and go to step 4. Otherwise go to step 4 without changing the bound.

Step 4 New Grid Points

If $L(\hat{x},\pi) < \sigma$ then use $\hat{x}$ to generate a new column for P_λ^t. Place the new P_λ^t on the list and go to step 2.

If $L(\hat{x},\pi) \geq \sigma$ then go to step 5.

Step 5 Optimality Test

Compute x^* from λ using (4). If $g_i(x^*) \leq 0$, $i=1,\ldots, m$, and if $f(x^*) < F^o$ then replace F^o with $f(x^*)$ and let x^* be the new incumbent solution.

If a) $f(x^*) \leq Z$ and b) $g_i(x^*) \leq 0 \quad \forall i$, then x^* is global optimal for the NLP subproblem over $x \varepsilon C^t$. Go to step 2.

If a) or b) is violated, go to step 6.

Step 6 Branch

Use x^* to generate a new column (nonbasic) for P_λ^t. Select a coordinate x_j^* $j=1,\ldots n$ of x^* (see discussion in section D.) and branch creating two new subproblems

a) P_λ^t restricted to $x_j \le x_j^*$ (include only those columns for which $(x_k)_j \le x_j^*$) and

b) P_λ^t restricted to $x_j \ge x_j^*$ (include only those columns for which $(x_k)_j \ge x_j^*$)

Compute a bound L^t for each of these subproblems and place both on the list. Go to step 2.

D. Branch and Problem Selection Rules

Two aspects of the method remain to be described: the rule for selecting the next subproblem P_λ^t to examine in step 2 and the rule for selecting the component x_j^* of x^* on which to branch in step 6. In this section we show that penalty calculations can be used for these decisions. The penalties used were originally considered in the context of integer programming and have been applied to separable nonconvex optimization using special ordered sets in [5]. In fact, neither separability nor the ordered set property is necessary as we shall show.

Consider the following linear program

$$\begin{aligned} &\min cx \\ \text{st} \quad &Ax = b \\ &x \ge 0 \end{aligned} \tag{11}$$

with optimal basis B and optimal solution

$$x_B = B^{-1}b \; ; \quad x_N = 0 \tag{12}$$

Suppose we have available the optimal simplex tableau containing the transformed constraint matrix

$$T = B^{-1}A \tag{13}$$

and also the row of reduced cost coefficients

$$C' = C - C_B B^{-1} A \tag{14}$$

A standard result of post optimality analysis is that if we force a basic variable x_{B_i} to zero, and let the remaining variables adjust optimally, then a first order approximation to the resulting objective function change is

$$Q_i = x_{B_i} \left(\min_{\substack{j \text{ nonbasic} \\ t_{ij}>0}} \left\{ \frac{C'_j}{t_{ij}} \right\} \right) . \tag{15}$$

This approximation is called the "penalty" for forcing x_{B_i} to zero. It gives the exact change if no basis changes occur before x_{B_i} reaches zero, and is otherwise an under-estimate.

In the context of branch selection for P_λ in step 6 of the algorithm we wish to compute a penalty for each component x_j for the two resulting subproblems. Each subproblem involves dropping several of the current vectors x_k, or equivalently forcing the corresponding variables λ_k to zero.

Suppose S is a set of variables which we want to force to zero if basic or maintain at zero if nonbasic. Then the penalty for this action is

$$Q_S = \max_{x_{B_i} \varepsilon S} \left\{ x_{B_i} \left[\min_{\substack{j \text{ nonbasic} \\ x_j \notin S \\ t_{ij}>0}} \left(\frac{C_j'}{t_{ij}} \right) \right] \right\} \tag{16}$$

In the context of P_λ, then, the appropriate penalties are obtained from (16) by setting S to be

$$S_j^+ = \{\lambda_k \mid (x_k)_j < x_j^*\} \tag{17}$$

for subproblem b) of step 6 and

$$S_j^- = \{\lambda_k \mid (x_k)_j > x_j^*\} \tag{18}$$

for subproblem a) for each $j=1,\ldots, n$.

Intuitively a subproblem with a large penalty is unlikely to contain the global optimal solution to NLP. Thus branch selection in step 6 can be performed so that one of the two resulting subproblems is most likely to contain the solution by selecting j to satisfy

$$\min_j \quad Q_j^{\pm} \tag{19}$$

or possibly

$$\min_j \mid Q_j^+ - Q_j^- \mid \tag{20}$$

In either case the choice of the next P_λ to work on in step 2 can then be the highly likely subproblem, placing the unlikely candidate on the list and hoping that it will be fathomed before it has to be solved.

It should be emphasized that in the nonconvex case the penalties yield only a guide and not guaranteed bounds on the new subproblems. Thus they should not be used to infer new bounds L_t on the resulting subproblems after a branch.

As in all branch and bound procedures the actual choices of subproblem and branch selection rules should be governed by computational experience.

REFERENCES

[1] Hartman, J. K. "Extension of the Grid Linearization Algorithm for Convex Optimization to Nonconvex Nonlinear Programs," Naval Postgraduate School, Tech. Report NPS-55Hh75071, July 1975.

[2] Hartman, J. K. "ε-Optimality for a Nonconvex Optimization Method," Naval Postgraduate School, Tech. Report NPS-55Hh75101, Oct. 1975.

[3] Hartman, J. K. "Some Experiments in Global Optimization", Naval Research Logistics Quarterly, Sept. 1973.

[4] Dantzig, G. B., Linear Programming and Extensions, Princeton, 1963.

[5] Beale, E.M.L. and J. A. Tomlin, "Special Facilities in a General Mathematical Programming System for Non Convex Problems using Ordered Sets of Variables," in Proc. 5th International Conference on Operational Research, 1969, John Wiley, New York.

INITIAL DISTRIBUTION LIST

	No. of Copies
Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
Dean of Research Code 023 Naval Postgraduate School Monterey, California 93940	1
Library, Code 0212 Naval Postgraduate School Monterey, Calif. 93940	2
Dr. Thomas Varley Office of Naval Research Arlington, Virginia 22217	2
Library, Code 55 Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, Calif. 93940	5
Professor Donald P. Gaver Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, Calif. 93940	1
Professor James K. Hartman Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, Calif. 93940	20

U170602

DUDLEY KNOX LIBRARY - RESEARCH REPORTS
5 6853 01071175 7

NPS